

# RMarkdown Demo 1

Alina Zeng

22/05/2021

**Download the datasets for this example script from:**

<https://github.com/ourcodingclub/CC3-DataManip>

**Install and load the relevant packages** \_\_\_\_\_

```
library(dplyr) # an excellent data manipulation package
```

```
##
```

```
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
```

```
##
```

```
##   filter, lag
```

```
## The following objects are masked from 'package:base':
```

```
##
```

```
##   intersect, setdiff, setequal, union
```

```
library(tidyr) # a package to format your data
```

```
library(pander) #to create pretty tables
```

```
library(knitr)
```

**Set your working directory to the folder where you have downloaded the datasets**

```
setwd("C:/Users/alina/Documents/git/Git_Tutorials/CC_course_stream2/05_Markdown")
```

**Import data** \_\_\_\_\_

```
elongation <- read.csv("EmpetrumElongation.csv", header = T) # stem elongation measurements on crowberry  
germination <- read.csv("Germination.csv", sep = ";") # germination of seeds subjected to toxic solution
```

## Tidying the data

---

Putting the data into long format using `gather()`

```
elongation_long <- gather(elongation, Year, Length, c(X2007, X2008, X2009, X2010, X2011, X2012))  
#gather() works like this: data, key, value, columns to gather. Here we want the lengths (value) to b  
head(elongation_long)
```

```
##   Zone Indiv  Year Length  
## 1    2   373 X2007    5.1  
## 2    2   379 X2007    8.1  
## 3    2   383 X2007    9.3  
## 4    2   389 X2007   15.0  
## 5    2   390 X2007    3.5  
## 6    2   395 X2007    6.1
```

## Investigating the data

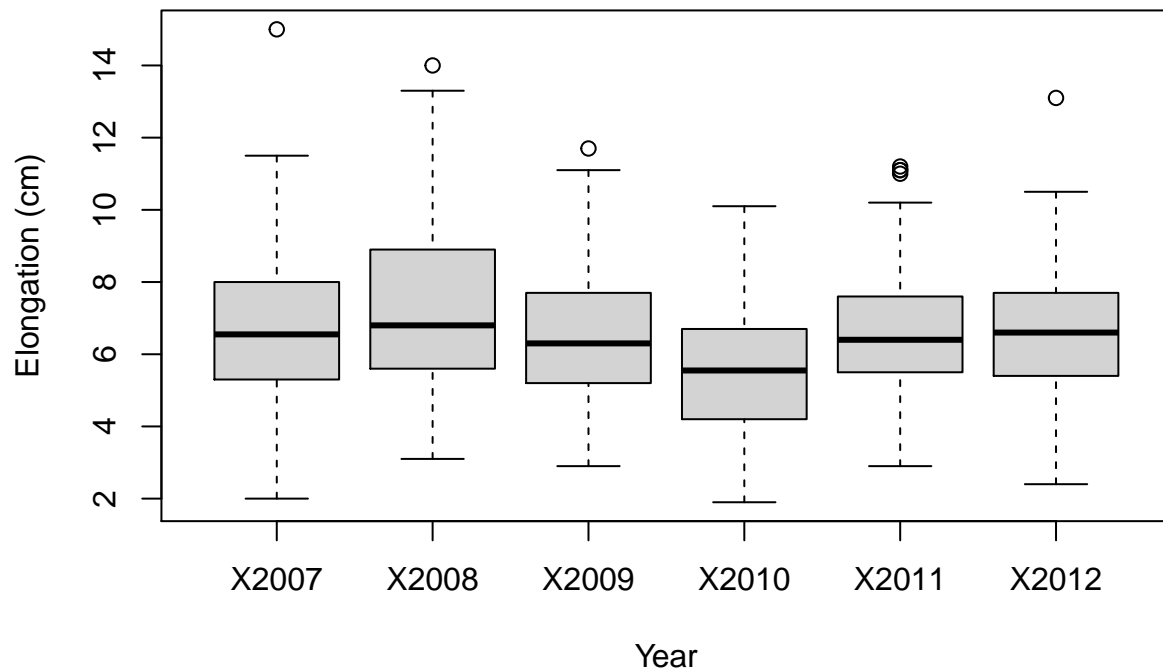
---

Create a boxplot of 'elongation\_long' to visualise elongation for each year.

This set of boxplots can be added to your R Markdown document by putting the code in a code chunk

```
boxplot(Length ~ Year,  
        data = elongation_long,  
        xlab = "Year",  
        ylab = "Elongation (cm)",  
        main = "Annual growth of Empetrum hermaphroditum")
```

## Annual growth of *Empetrum hermaphroditum*



Use `filter()` to keep only the rows of `germination` for `speciesSR`

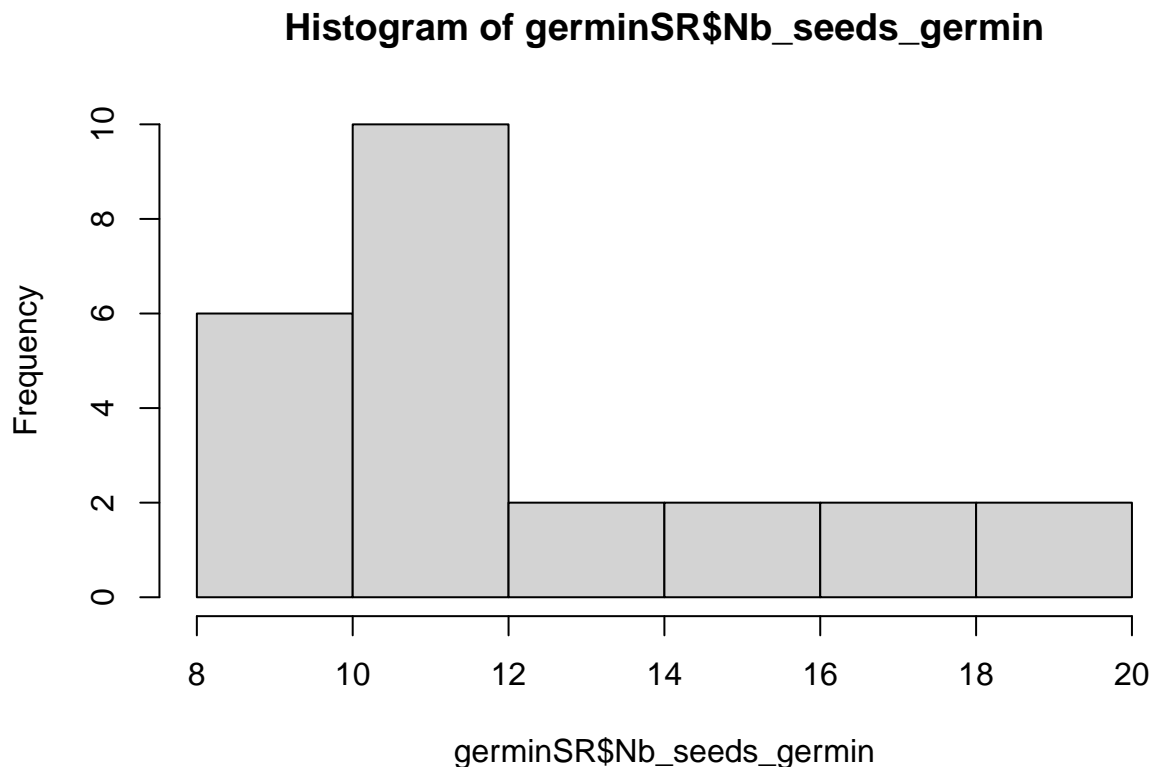
```
germinSR <- filter(germination, Species == 'SR')
```

Let's have a look at the distribution of germination across SR

This histogram can be added to your R Markdown document by simply putting the code in a code chunk

Try adding some plain text to your R markdown document to explain the histogram

```
hist(germinSR$Nb_seeds_germin, breaks = 8)
```



Use `mutate()` to create a new column of the germination percentage using the total number of seeds and the number of seeds that germinated

```
germin_percent <- mutate(germination, Percent = Nb_seeds_germin / Nb_seeds_tot * 100)
```

Use a pipe to get a table of summary statistics for each Seed type

```
germin_summ <- germin_percent %>%  
  group_by(Species) %>%  
  summarise("Mean germination per" = mean(Nb_seeds_germin), "Max germination per" = max(Nb_seeds_germin))
```

Make a table of 'germin\_summ' in your R markdown document using `pander()`, the instructions can be found in the tutorial

```
pander(germin_summ)
```

Species	Mean germination per	Max germination per	Min germination per
SP	17.08	23	13
SR	12.5	20	8